

小木球笔记——朝闻道，夕如何？

——一枚酱油党在研读营上的听雷拔草之旅

引子

关键词：闻道，祛魅，工具的局限。

在古北的最后一天下午，回程的车因故推到晚上，又多了几小时与这些人混迹的时间。我折回这座优美的仿古小镇去寻找大部队，感觉像磕了药，心中有一种极其放松和超然的欣悦感，看着满眼的游人、红叶和水镇秋景，似乎自己的身体瞬间就会漂浮起来。

科幻作家大刘有个短篇叫《朝闻道》，讲外星人挑选了一拨地球物理学家，解答他们关于宇宙的一切疑问，前提是，听完当晚他们就要死掉。有朋友曾问过我感觉最幸福的时刻是什么？大约就是这种时刻吧，这种豁然开朗，大脑打开的时刻。至于死不死的，先吃了这顿饭再说吧：目。

我无意刻意拔高这次研读营的规格和前沿性，这应该由物理学家来评价。仅是对于我这样一个平时只能靠科幻作品理解高等物理知识的生物狗来说，如小鸭听雷，有幸能旁听到这群科研一线的理论物理大牲口们的对撞，尤其是解开了我长久以来的许多科幻的疑惑，完成了对一些神秘事物的一次“祛魅”，似乎抵达了现有认识的边界。

同时，对于我这样一个业余的复杂系统粉，对复杂科学的各种数理和计算机工具，研究系统的方法，一直充满着敬畏感和神秘感。在这个学科的门口卧底数年，也想偷门大炮回去打鸟，跟踪到今天，终于触摸到复杂系统研究的极限，知道数理计领域最前沿的思想能够解决什么问题，试图解决什么问题。终于意识到这些工具对于生命问题的解决仍然存在巨大的局限性和鸿沟，而系统生物学可能以及甚至已经在给复杂系统研究带来新的机会。

一、一句话简介

关键词：复杂网络的双曲嵌入、重整化、量子比特、演生时空

本次研读营的主题是《网络、几何与机器学习》，首先，从自己理解的角度，大概描述一下这五天里每个人讲的核心问题。

第一天：【张江】回顾了复杂网络的研究历史，发现无尺度网络和小世界网络均和实证有冲突，其中的高聚集效应可以网络粗粒化和重整化之后便于研究，但在这个尺度变换时，有什么是不变的呢？现在发现神经网络可能可以发现复杂网络的特性，同时考虑流动性，并解决实证的问题。**【尤亦庄】**介绍双曲几何的三个模型：扩展的庞加莱模型、庞加莱圆盘以及半平面模型，如何将复杂网络嵌入到双曲模型中并求解测地线方程，以及通过重整化获取网络特征。**【李嫣然】**讲复杂网络如何嵌入神经网络的问题。

第二天：【张潘】对于统计物理的网络，其算法难度和正反两个问题。正问题是将所有网络

问题嵌入到玻尔兹曼分布中，通过模型参数得到宏观统计量（如hopfield模型），并给出了数学推导过程；反问题是知道外界输入，学习到网络的特征模型，并给出了如何重整化和嵌入神经网络的数学方法。

第三天：【尤亦庄】讲第二次量子革命（第五次物理学革命）背景下如何理解空间的起源与网络的关系。物质由量子比特中演生，退相干即与环境耦合，信息丢失，扩散到全局，即为熵的起源和时间的方向性，而空间是整个量子态的纠缠结构。量子纠缠的非局域性由虫洞相连，实际是局域的。量子信息存在于量子比特的关系中，可以用张量网络来描述，纠缠态通过哈密顿量求解。

第四天：【吴令飞】大数据与计算社会学的发展现状，由面对对象的科学转为面向关系的科学，沿用物理学外场逻辑的一些学说。应为数据和研究问题驱动，而非模型驱动研究。进一步研究网络重整化问题。**【新人专场】**甄慧玲、董磊、傅渥成、章彦博、程嵩、罗秀哲、曾培等介绍各自的研究。**【老人专场】**王雄、苑明理、王成军、小木球介绍自己感兴趣的方向。

第五天：【韩峰】讲熵、量子与区块链的关系。经典物理给不出时间箭头，熵最大系统是人类观测到的局部，不会走向热寂。对量子做低复杂度的测量，不能带来系统全部信息，产生时间的方向性。麦克斯韦妖解决了朗道原理的单分子热机模型，区块链就是分布式的麦克斯韦妖智能系统，通过全网基础协议给支付盖时间戳，剔除重复交易。重整化群就等于演生，深度学习就是去找不变点即收敛熵，量子重整化群可能成为未来区块链研究的普适性数学工具。**【文教授】**介绍分形的研究历史，嵌入双曲后的度量，以及重整化。

二、奇异的跨界：复杂网络的新生

关键词：庞加莱圆盘、测地线、重整化、拓扑旋

1，网络嵌入几何

这部分和机器学习，其实是本次研读营最核心和精华的地方，可惜我数学不好，几乎听不懂，只能抓住几个关键词，囫圇猜测一下其中的意思。

机器学习和复杂网络为什么会有关系呢？我的理解，是因为复杂事物和复杂网络是机器学习最重要的对象。深度学习的目的，就是要让人工智能取代人，去学习复杂网络表现出来的特征，从而能够模拟或者预测复杂系统的行为。

如果复杂网络能够嵌入双曲几何，并用拓扑和双曲相关的数理工具进行解析，那么就可以使网络变得可能被理解。庞加莱圆盘，就是可以内嵌复杂网络的双曲模型。测地线，或者说哈密顿量，其实就是这个拓扑结构中的最短路径。求解测地线，可以说是对这个网络的动力结构进行描述。

重整化，是对复杂网络进行描述的数学方法，简单说就是按照一定规则将网络中的数个节点看做一团，抓取其中的一些属性，同时忽略另外一些属性。这样，就把难以解决的大问题，变成一层层可以解决的小问题。好的重整化角度，能够提取出复杂网络非常重要的特点和属性，这里尺度和角度的选取很重要。

有趣的是，在自然的复杂系统演生过程中，又有粗粒化这样一种现象和重整化相对应。如果说重整化是人们理解网络结构时人为的一种划分模式，那么粗粒化则是网络结构中自然发生的聚类现象和网络层次性的源头。如何用重整化的角度来研究粗粒化，或者如何通过粗粒化找到重

整化的规律，可能是探讨网络层次性发生的重要角度。

2，终极挑战：复杂网络统计属性与解析属性的统一之路

第四天下午是新人时间，晚上是老人时间。没讲课和报告的参与者也有机会向大家介绍自己的工作和想法。其中苑明理的短小的报告给我留下极深的印象，当然也引起了大家足够热烈的关注和讨论。这可能只是研读营里一个很小的插曲，也可能今后是一个极大的亮点。

老苑是中文维基的创始者，也是彩云重要的算法工程师和合伙人，还有一个著名的瓦克星教育计划，但同时，他也是一位对本质问题有着深度思考的数学家。这一次，他带来了一个不起眼的小公式，当然公式我是看不懂了，但知道这个公式的妙处在于，将机器学习得到的特征函数与描述底层的微分方程联系到了一起。

要知道机器学习出现以来，便面临了巨大的质疑和争论。在去年网络智库的一次会议现场，肖达讲解神经网络和深度学习方法后，即有一位老师强烈质疑，认为深度学习如同中医，是“知其然而不知其所以然”，不能获得对系统机理的丝毫理解。

不能说这个质疑毫无道理，毕竟神经网络对复杂系统特征的学习，获得的还是系统的统计性质，而对这些统计斑图是从何而来并不能给出解释。但是之前从底层Agents相互作用开始讨论的解析方法，则对数据量巨大、算法难度极高的多主体网络十分吃力。统计性质与底层机理的统一，可能是复杂系统研究的终极挑战，或者说是整体论和机体论的终极解谜。

这个公式的提出，是一个开始，标志着对这一统一的征程已经迈开了步伐。但仅仅将两套系统的结果联系在一起还是远远不够的，如何找到将机理和统计性质进行直接关联的方法理论模型，可能是下一步更重要的事。

3，拓扑旋：一个小小的脑洞

拓扑旋并不是研读营上听到的，而是偶然的一个机会在程嵩的朋友圈里看到，他向大家科普拓扑学时举的一个例子。物理学家关注系统中所有恒定的现象，旋或者洞，是双曲几何中非常重要的拓扑不变量。

大家知道复杂网络的流动性，或者说其动力学机制是比较难描摹的，而另一方面，动力学，也是揭示复杂网络演化机理和起因的重要线索。之前所说的测地线，可能从结构上解释这一网络为何与另一网络不同，某一网络的静态结构是如何的。但如何从机理角度解释网络演生的动力，则尚不清楚。

我们知道自指环（或者说自复制子）是一个稳定的结构，拓扑学可以告诉你为什么环状最稳定，因为它构成了一个拓扑不变量。艾根的超循环理论告诉我们，生命系统的演生，起源于自复制子的自催化循环。一旦系统的初始自指环建立，会以具有方向性的漩涡状结构，将系统周边的其它物质逐渐纳入，并不断自加强。在双曲几何中，可以看到这个圆盘的曲率不断加深。在系统足够大时，由于结构的复杂化，原初的自指环可能难以找寻。但如果假设一个复杂系统一定有一套起始循环，那么其动力系统也一定是具有方向的大循环，由此也可以找到嵌入双曲面的起始“旋”。我猜测，这，也许是连接机理与统计，连接静态网络与动力学机制的可能入手点。

三、演生宇宙：科幻粉的终结

关键词：演生的宇宙，多重宇宙，时间逆转（纯YY，数理专业请跳过。）

作为一枚只有中学和科幻世界物理水平的资深科幻粉，暗物质、多重宇宙、时间逆转……这些词既听不懂，又让人着迷。幽冥之中，越深越暗越美丽。一方面给人种下了对物理学的向往，另一方面随这些词语埋下的神秘感也成为了获取知识的障碍。

温文瘦弱的大男孩尤亦庄由于和文小刚教授合作研究基于拓扑序的量子多体理论，因此成为研读中被追捧的大牛。他们讨论的那些烧脑问题我都听不懂，作为物理小白和科幻粉，我也（替大家#&误）问了几个心中深藏已久的脑洞问题，这些明确肯定的回答让我非常兴奋，以后科幻与我无关了哎呦喂：）。

1，宇宙演化与对称性破缺

之前，经常听到物理学家用双螺旋的手性，以及心脏的位置来科普对称性破缺，以为是一个比喻，还是一个不太恰当的比喻。因为之前接触到的对称性破缺，讲的总是微观量子世界的事情，与宏观的不对称性似乎不是一回事。另一个比喻，以人与人之间的相互关系来表述量子纠缠的涨落及长程，也有很多违和。

我们都知道生命界的许多不对称性是自然选择的结果。最简单的例子就是DNA双螺旋的手右性，人工可以构造出左、右手DNA甚至三螺旋，但在地球上所有已知生命的体内，都只有右手双螺旋一种可以被利用，为什么呢？这是因为作为所有生命共同祖先的那个（群）细胞，碰巧是以右手螺旋保存遗传信息的，而因为各种原因（选取右手螺旋本身可能不是原因），左手螺旋被演化淘汰掉了。同样也是偶然，右侧心脏的脊椎动物祖先被淘汰掉了。如果微观的对称性破缺是物理世界的现象，那么怎么能跟以上演化和适应扯上关系呢？

尤大大的回答是：在我们可以观测到的宇宙起源初始，反物质和正物质是均匀存在的，我们现在看到的正物质世界，也是在和反物质的较量及选择中胜出的结果，这种对称性破缺，和右手螺旋、左侧心脏的胜出是一样的。

宇宙起源于无时无处不在的量子纠缠和退相干，实际上退相干和对称性破缺就可以看成是一种选择的过程，从无数可能的状态中选择塌缩为确定的状态，从而不断向外扩散，一路演化下来。由此，生命的诞生和演化，虽然有其独特和神奇之处，但也其实只是整个宇宙演化路上的一个延续。选择无处不在，胜负不一定由于优劣，更多是偶然。一旦被选择，便固化进了永恒的存在。宇宙的演化理论自数学家和物理学家Philip Anderson以来，逐渐被物理学界所接受并得到了很好的发展。将宇宙的演化与生命的起源在时序上连接起来，它们共享同一套底层原理，包括量子退相干产生结构及其延续，以及偶然性选择在对称性破缺中的作用，以此角度重新审视和完善生物学的演化理论，是非常必要和自然的。

2，暗物质去哪儿了，有没有多重宇宙？

接着上面宇宙演化的问题。正物质与反物质相遇湮灭并放出巨大能量，那么会不会像有些科幻小说里写到的，在我们生活的正物质空间的同一空间中，还有一个反物质空间，其时间箭头和我们的世界相反，仅只是两个世界无法交互信息？甚至于，在同一空间中，是否真实存在着多重的宇宙以及我们的无数可能性？

回答是：（在我们可以观测到的局部）宇宙演化过程中，暗物质（反物质）在和正物质的较量中大部分已经湮灭，在我们共同的空间中，并不存在所谓另一个反向的宇宙，更不存在无数个同时发生的宇宙作为现在时间点的分支。这些都只是物理学的思想实验，并不是实际存在的情况。

3、时间不可逆是否可能有局部例外？

大家记得《你一生的故事》里面的情景：全局的时间信息被外星生物观察到并付诸其文字。这种幻想是否具有可能性？英国边缘生物学家谢尔德雷克在形态发生场假说里猜测后续事件的因果反过来影响之前的因果，这种脑洞型假设是否有依据？

上面讲到，时间箭头是由于人类观测者无法获得全局信息而产生的幻觉，因此除非观测到全局，否则无法看到时间的可逆性。但是当脑洞一个情景：如果虫洞是量子长程纠缠引起的空间形变，那么会不会有某种不平衡现象，即信息只扩散到了一部分结构而非全局，而这一部分信息传播的结构有其特殊的空间形变，从而可以通过这些结构示踪到信息的传导路径特征，从而推断出整个空间的信息？

回答是：在量子多体宇宙中，纠缠和去纠缠的涨落时刻发生，没有任何结构和纠缠能够与周围绝缘，信息只在特定结构中扩散是不可能存在的，因此也不可能出现能够观测到全局的外星生物，不可能在局部出现时间逆转。

如此，关于时间的科幻，就不要心存妄想了，洗洗睡足，可以起来关注演化和复杂系统中更实际的问题了。

四、生命复杂性：置之死地而复生

关键词：局限性、生长点

最后来说生命复杂性问题。生命复杂性一直是复杂科学试图探讨和解决的问题，并有了近一个世纪的努力。到目前来看，物理思维讨论生命问题仍有以下几个倾向：1) 过度简化，忽视生命现象的层次性；2) 形而上，重斑图轻机理；3) 将物理复杂系统与生命复杂系统混同，将社会复杂系统和物理复杂系统等同。

作为生命科学研究者，有时候不太喜欢物理学家的一些类比，将非常底层的物理现象与很高级的生物或社会现象做直接联系和比较。虽然这种思维方式是复杂科学起始时便具有和奠定下来的，但有时会忽略掉物理复杂系统和生命复杂系统以及基于生命的社会复杂系统之间的区别，同时忽略掉生命复杂系统本身的机理和层次。

比如希望用深度学习的方法来学习生命网络的结构，就要考虑讨论的是哪一层次的结构，更重要的是，在生命科学的技术中，是否可以或容易获得足够的训练数据可以用来学习和提取特征。对于目前来说，只有比较大的基于序列的生物信息学数据库，以及病理图像数据比较能方便这种应用。而对于生物网络数据，或者说各种组学数据，因其繁杂，是否已积累了足够的可学习数据，则需要更具体和详细的探讨。

因此，虽然本次参会我的主要任务是会务筹备、打字和打酱油，但为了在物理大牛堆里刷一下生物狗的微弱存在感，除了提问小白问题，还是做了几件小事：1) 尽量以物理生能理解的角度向大家介绍生物系统的复杂层次。系统生物学中，对生命系统的层次性有很好的表述和研究，我尝试以物理学者能理解的角度进行了简单的介绍，尤其提到了表观遗传学角度对网络调控体系探讨的重要性；2) 与傅渥成、王雄探讨了生命起源中对形成自复制子的分子有什么抽象特点，以及为什么复制-翻译循环的建立是起源问题的关键；3) 向傅渥成的请教中，了解到建模方法对于生命网络和生命现象解释和预测的难度；4) 收获了2个小问题：A，（王雄）为什么核桃像大脑，大脑像核桃？B，（傅渥成）为什么生命信息分子是线性排列而非分形结

构？非常喜欢这两个小问题，现在闲暇时也常会在脑中思考把玩。见此文有兴趣的朋友也可以一起探讨。

穷尽其端也尚未搜罗到好的武器，也算是体会到撞到了“十年寻梦路，一头撞南墙”的感觉。但意识到复杂研究工具对于生物学研究的局限性，了解了方法的边界也是好的。

另一方面，我发现，从面向对象到面向关系，复杂科学迈向网络研究的方向，但目前复杂网络研究的主要对象还是单层网络的结构，而对于生命这种多层网络相互嵌套、互作的结构来说，网络之间的调控和通信如何描述，可能是今后一个很大的课题。对于此问题，物理学尚未涉及，但现有的系统生物学和抗肿瘤免疫研究似乎已经有了一些尝试和探索。生命网络无疑是复杂性研究新的生长点，在接下来的时间里，系统生物学也许会进一步成为复杂系统更深入发展的不可或缺的关键战场。先挖个坑，待以后有机会专门展开吧。

六、和风微熏处，交杂呈异彩

听凯风基金会的老师说，“凯风”出自诗经，用以比喻如母爱般温暖不易察觉的南风。集智和凯风能够联姻，既有偶然，似乎又出于一种必然。这个必然基于这两个组织都有同样一种风格，一种如凯风般温暖不易察觉的行为方式，以及一种既疏离又积极的气质。疏离在于，在多年来学界也好、社会也罢，各种浮躁的、商业为先的、急功近利的气氛下，难得会有凯风和集智这样的组织，刻意与大环境保持距离，埋下头来做别人不肯做的慢功夫。积极在于，虽然疏离，却又时刻关注时代，走在最前端人迹罕至的领域去探索开拓，不经意间就引领了潮流。

学科交叉交流一直是集智的特色与主线。能够举办这样一个高规格高密度的科学前沿研读营，这种跨学科交流的规格和密度，让每一位参与者兴奋不已。Jake说，这次活动，是集智所有以往活动中，他思考得最少，准备得最少，但效果最好，收获最大的一次。当然得益于凯风会务人员专业的组织力，更重要的是，这是一个水到渠成自然而然的事情。集智科学家群体常年来深厚的积累和探讨，在科学理想上的坚持、探索和不妥协，都是渠中之水，不引自来。

中间的拍照时间，美女摄影师Amanda讲了一个对比，她说经常给文艺界拍照，需要协调抢镜现象，而给这一大班理科生拍照，却要不停地把大家拽到前面来。在破冰会上必须组织挨个发言，否则似乎都不知道怎么去和别人打招呼。这么一群谢耳朵聚在一起，大家都似乎很沉默拘谨，不善社交，看起来灰溜溜的。但一旦把其中任何一个放到台上开讲，或者相互交流他们感兴趣的话题，这些人就会像是被按了一个神秘开关，切换到一种“点亮”状态，眉飞色舞，滔滔不绝，一旦讲完，又回到“熄灭”状态了。

这就是谢耳朵的行为方式，实际上，营中的每一位参与者在交流碰撞中都很激动，这是毋庸置疑的，新的好友、闺蜜在这群体中产生，当然更多的还有新的启发和合作点子。最近在看维纳的《控制论》，前言中介绍了当时维纳、香农、冯诺依曼、费希尔，还有神经网络先驱Pitts & McCulloch、医学家Rosenblueth等人经常在一起进行跨学科研讨，当然结果不仅仅是开创了控制论这个学科，后续这些大牛们的深远影响和鼎鼎大名大家也都有所闻。这样的学术交叉，也可以说是Jake大大发起集智俱乐部的初衷。

但是，其实凯风基金会还有更大的野心：就是希望在这基础上，能够进一步使得理科的前沿进展和人文学科建立交流，对科技进步对社会的改变等问题有更深入地探讨，以及建立对整个社会更科学和深刻的认识。我也在期待这一水平的交流早一天到来。

尾声

最后，回到主题。朝闻道，夕如何呢？全局信息不可能被获得，求道无止境。研读营最赞的收获是，大家不仅只是交换了不同学科的信息，而是更加明确了方向，脑力激荡的满足感之后，脚下的路会更加清晰。

这就是我的体会。

小木球 2016.11.16